

BEST AVAILABLE TRANSLATION FOR JAPANESE PUBLICATION NO. 2003-504674  
(APRIL 2, 2003).

In a speech recognition device (1) comprising receiving means (36) for receiving voice information (AI) uttered by a speaker and including speech coefficient memory means (38) for storing a speech coefficient indicator (SKI, PRI, SMI, WI) and including speech recognition means (42) which are arranged for recognizing text information (RTI) corresponding to the received voice information (AI) by evaluating the voice information (AI) and the speech coefficient indicator (SKI, PRI, SMI, WI), and including correction means (49) for correcting the recognized text information (RTI) and for producing corrected text information (CTI) and including adjusting means (50) for adjusting the stored speech coefficient indicator (SKI, PRI, SMI, WI) by evaluating at least the corrected text information (CTI), text comparing means (52) are provided for comparing the recognized text information (RTI); with the corrected text information (CTI) and for determining at least a correspondence indicator (CI) and the adjusting means (50) are provided for adjusting the stored speech coefficient indicator (SKI, PRI, SMI, WI) by evaluating only one of such text parts (P2) of the corrected text information (CTI) whose correspondence indicator (CI) has a minimum value (MW).

일본 공개특허공보 특개2003-504674호(2003.02.04.) 1부.

(19) 日本国特許庁 (J P)

(12) 公表特許公報 (A)

(11) 特許出願公表番号  
特表2003-504674  
(P2003-504674A)

(43) 公表日 平成15年2月4日(2003.2.4)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	データシート*(参考)
G 1 0 L 15/00		G 1 0 L 3/00	5 5 1 B 5 D 0 1 5
15/06			5 6 1 E
15/18			5 2 1 H
15/22			5 6 1 C
			5 3 7 C
審査請求 未請求 予備審査請求 未請求(全 37 頁)			
(21) 出願番号	特願2001-509020(P2001-509020)	(71) 出願人	コーニンクレッカ フィリップス エレクトロニクス エヌ ヴィ Koninklijke Philips Electronics N. V. オランダ国 5621 ベーアー アインドーフェン フルーネヴァウツウェグ 1, 5621 BA Eindhoven, The Netherlands
(86) (22) 出願日	平成12年6月30日(2000.6.30)	(72) 発明者	バルトシク ヘインリッヒ オランダ国 5656 アー アー アインドーフェン プロフホルストラン 6
(85) 翻訳文提出日	平成13年3月8日(2001.3.8)	(74) 代理人	弁理士 沢田 雅男
(86) 国際出願番号	PCT/EP00/06167		
(87) 国際公開番号	WO01/004874		
(87) 国際公開日	平成13年1月18日(2001.1.18)		
(31) 優先権主張番号	9 9 8 9 0 2 3 2. 4		
(32) 優先日	平成11年7月8日(1999.7.8)		
(33) 優先権主張国	欧州特許庁 (EP)		
(81) 指定国	EP(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, I T, LU, MC, NL, PT, SE), JP		

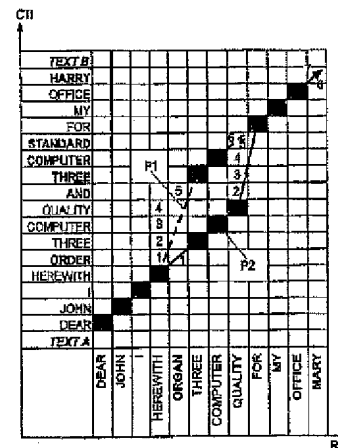
最終頁に続く

(54) 【発明の名称】 スピーチ認識装置とテキスト比較手段

(57) 【要約】

【課題】 受信されたボイス情報と十分に関係する訂正されたテキスト情報のみが、スピーチ係数インジケータの調整に使用される、スピーチ認識装置とスピーチ認識方法を提供すること。

【解決手段】 スピーチ認識装置(1)であって、話し手によって発声されたボイス情報(AI)を受信するための受信手段(36)を有し、スピーチ係数インジケータ(SKI, PRI, SMI, WI)を格納するためのスピーチ係数メモリ手段(38)を有し、ボイス情報(AI)とスピーチ係数インジケータ(SKI, PRI, SMI, WI)を評価することによって、受信されたボイス情報(AI)に一致するテキスト情報を認識するように構成されているスピーチ認識手段(42)を有し、認識されたテキスト情報(RTI)を訂正し、訂正されたテキスト情報(CTI)を生成するための訂正手段(49)を有し、少なくとも訂正されたテキスト情報(CTI)を評価することによって、格納されたスピーチ係数インジケータ(SKI, PRI, SMI, WI)を調整するための調整手段(50)を有する、スピーチ認識装置において、テキスト比較手段(52)が、認識されたテキスト情報(RTI)と訂正されたテ



【特許請求の範囲】

【請求項1】

スピーチ認識装置であって、  
話し手によって発声されたボイス情報を受信するための受信手段と、  
スピーチ係数インジケータを格納するためのスピーチ係数メモリ手段と、  
前記ボイス情報と前記スピーチ係数インジケータを評価することによって、前記受信されたボイス情報に一致するテキスト情報を認識するように構成されているスピーチ認識手段と、  
前記認識されたテキスト情報を訂正し、かつ訂正されたテキスト情報を生成するための訂正手段と  
少なくとも前記訂正されたテキスト情報を評価することによって、前記格納されたスピーチ係数インジケータを調整するための調整手段を含むスピーチ認識装置において、  
前記認識されたテキスト情報と前記訂正されたテキスト情報を比較し、少なくとも1つの一致インジケータを決定するためのテキスト比較手段が設けられていて、かつ

前記調整手段が、一致インジケータが最小値を有する前記訂正されたテキスト情報の1つのテキストパートのみを評価することによって、前記格納されたスピーチ係数インジケータを調整するように構成されていることを特徴とするスピーチ認識装置。

【請求項2】

前記テキスト比較手段が、評価された一致インジケータが前記最大の合計一致インジケータを有する前記訂正されたテキスト情報のテキストパートのシーケンスを決定するように構成されていて、かつ前記調整手段が、前記訂正されたテキスト情報の少なくとも前記決定されたテキストパートのシーケンスを評価することによって、前記格納されたスピーチ係数インジケータを調整するように構成されていることを特徴とする請求項1に記載のスピーチ認識装置。

【請求項3】

音声モデルデータをスピーチ係数インジケータとして前記スピーチ係数メモリ

手段の中に格納することが可能であって、このスピーチ係数インジケータが、ボイス情報内の少なくとも第一ワードに続く第二ワードの確率を示すことを特徴とする請求項1に記載のスピーチ認識装置。

【請求項4】

コンテキスト情報をスピーチ係数インジケータとして前記スピーチ係数メモリ手段に格納することが可能であって、このスピーチ係数インジケータが、ボイス情報の中の前記スピーチ認識手段によって認識できるワードすべてを含むことを特徴とする請求項1に記載のスピーチ認識装置。

【請求項5】

言葉参照情報をスピーチ係数インジケータとして前記スピーチ係数メモリ手段に格納することが可能であって、このスピーチ係数インジケータが、話し手の発音の特性に合わせた前記スピーチ認識手段の調整を可能にすることを特徴とする請求項1に記載のスピーチ認識装置。

【請求項6】

ボイス情報に一致するテキスト情報を認識するスピーチ認識方法であって、前記方法が、

ボイス情報を受信するステップと、

前記受信したボイス情報と格納されているスピーチ係数インジケータを評価し、かつ認識されたテキスト情報を生成するステップと、

認識されたテキスト情報を訂正し、かつ訂正されたテキスト情報を生成するステップと、

少なくとも前記訂正されたテキスト情報を評価することによって、前記格納されたスピーチ係数インジケータを調整するステップと

を有するスピーチ認識方法において、

前記認識されたテキスト情報を、前記訂正されたテキスト情報と比較し、かつ一致インジケータを決定し、かつ

前記格納されたスピーチ係数インジケータを調整するために、一致インジケータが最小値を有する、前記訂正されたテキスト情報の1つのテキストパートのみを使用することを特徴とするスピーチ認識方法。

【請求項7】

一致インジケータが前記最大の合計一致インジケータを有する、前記訂正されたテキスト情報のテキストパートのシーケンスを決定し、かつ、前記格納されたスピーチ係数インジケータを調整するために、少なくとも、前記訂正されたテキスト情報の前記決定されたテキストパートのシーケンスを使用することを特徴とする請求項6に記載のスピーチ認識方法。

【発明の詳細な説明】

【０００１】

【発明が属する技術分野】

本発明は、請求項１の於て書きに記載のスピーチ認識装置に関し、また、請求項６の於て書きに記載のスピーチ認識方法にも関する。

【０００２】

【従来技術】

このようなスピーチ認識装置とこのようなスピーチ認識方法は、US 5,031,113 から公知である。この公知のスピーチ認識装置は、ディクテーションにおいて話し手によって発声されるスピーチ伝達情報を受信および格納するためのマイクロフォンとオーディオメモリによって形成される受信手段を含む。

【０００３】

公知のスピーチ認識装置は、さらに、スピーチ認識方法の実行に必要なスピーチ係数インジケータが格納されるスピーチ係数メモリ手段を含む。このスピーチ係数インジケータは、コンテキスト情報、音声モデルデータ、音素参照情報を含む。コンテキスト情報は、スピーチ認識装置によって認識できるワードすべてを含み、音声モデルデータは、ボイス情報の中のコンテキスト情報のワードのシーケンスの確率を含み、音素参照情報は、ワードポーション（word portion）（音素）が話し手によってどのように発音されるかの情報を含む。

【０００４】

公知のスピーチ認識装置は、さらに、スピーチ認識方法の実行中に、スピーチ係数メモリ手段に格納されているスピーチ係数インジケータを評価することによって、受信されたボイス情報に一致するテキスト情報を認識し、そのテキスト情報を認識されたテキスト情報として送るように配置されたスピーチ認識手段を含む。認識されたテキスト情報は、モニタによって表示される。

【０００５】

テキスト処理プログラムとキーボードは、モニタによって表示された認識されたテキスト情報を訂正し、訂正されたテキスト情報としてモニタに再び表示する訂正手段を形成する。基本的に、使用者は、スピーチ認識処理時に誤って認識さ

れたテキスト情報のワードを、実際に話したワードに置き換え、その一方で、別の訂正も行う。別の訂正とは、例えば、住所など標準的なテキスト部分を例えば、挿入したり、ディクテーション時に話し忘れたテキスト部分を挿入する、あるいは認識されたテキスト情報のテキスト部分をキーボードによって入力されるテキスト情報に置換するなどである。

【0006】

公知のスピーチ認識装置は、前に撰って認識されたワードを、それ以降にスピーチ認識処理において正しく認識するために、話し手と言語に対するより良好な調整によってスピーチ係数インジケータを調整するための調整手段を含む。コンテキスト情報と音声モデルデータを調整するため、訂正されたテキスト情報が評価され、音素参照情報を調整するため、オーディオメモリに格納されているボイス情報も評価され、調整されたスピーチ係数インジケータを生成し、それをスピーチ係数メモリ手段に格納する。

【0007】

この公知のスピーチ認識装置とこの公知のスピーチ認識方法では、スピーチ係数インジケータの調整に使われる訂正されたテキスト情報に、ボイス情報とまったく関係ないテキスト部分も含まれてしまうことが欠点であることが判明している。スピーチ係数インジケータの調整にこのようなテキスト部分が使われると、調整の後に、スピーチ係数インジケータが、話し手と言語にとってより良好ではなく、より悪く調整されているということが起こりうる。

【0008】

【課題を解決するための手段】

本発明の目的は、受信されたボイス情報と十分に関係する訂正されたテキスト情報のみが、スピーチ係数インジケータの調整に使われる、スピーチ認識装置とスピーチ認識方法を提供することである。この目的は、請求項1の於て書きに記載のスピーチ認識装置において、請求項1の特徴記載部の方策によって、および請求項8の於て書きに記載のスピーチ認識方法において、請求項8の特徴記載部の方策によって、達成される。

【0009】

このことは、スピーチ係数インジケータの調整の前に、訂正されたテキスト情報に、大きく訂正されたテキストワード、またはまったく新しく挿入されたワードが含まれるか否かのテストが行われ、そのようなテキストワードがスピーチ係数インジケータの調整に使われないことを達成する。利点として、スピーチ係数インジケータが調整されるたびに、スピーチ認識装置とスピーチ認識方法の認識率がかなり向上する。

【0010】

請求項2と請求項7の方策によって、受信されたボイス情報または認識されたテキスト情報にそれぞれ十分に関係する認識されたテキスト情報のテキストワードが、テキストワードのシーケンスに連結される。最大の合計一致値を有する認識されたテキスト情報のテキストワードのシーケンスが、調整に使われる。このことは、最大的一致インジケータを有するテキストワード群の中のテキストワードも調整に使われ、このため、スピーチ係数インジケータの調整のたびに、スピーチ認識装置とスピーチ認識方法の認識率がさらに向上するという利点を提供する。

【0011】

請求項3、4、5の方策によって、スピーチ係数インジケータに含まれる情報すべてが非常に良好に調整される。

【0012】

図に表されている実施例の3つの応用例を参照しながら、以下に、本発明についてさらに説明する。ただし本発明は、この応用例に限定されない。

【0013】

【発明を実施するための形態】

図1は、スピーチ認識方法によるスピーチ認識プログラムが実行されるコンピュータ1を示し、このコンピュータ1は、二次的なスピーチ認識装置を有するディクテーションマシンを形成する。

【0014】

コンピュータ1には、手で持つことができる入力ユニットを形成するディクテーションマイクロフォン2と、足で作動させることができる入力ユニットを形成



するフットスイッチ8と、モニタ4と、キーボード5が接続されている。図2は、ディクテーションマイクロフォン2を、ブロック線図の形式でより詳細に示す。

【0015】

ディクテーションマイクロフォン2は、オーディオ受信手段を形成するマイクロフォン6を有し、使用者によってマイクロフォン6に発声されたスピーチ伝達情報を受信して、第一アナログオーディオ信号A1を送るように構成されている。ディクテーションマイクロフォン2は、さらに、第一アナログオーディオ信号A1を処理して、オーディオ情報A1としてデジタルオーディオデータを送るように構成されているオーディオ処理手段7を含む。

【0016】

オーディオ処理手段7は、第一増幅段8、アナログ-デジタルコンバータ段9、第二増幅段10、デジタル-アナログコンバータ段11を含む。第一アナログオーディオ信号A1は、第一増幅段8に与えることができ、増幅された第一アナログオーディオ信号は、アナログ-デジタルコンバータ段9に与えることができる。アナログ-デジタルコンバータ段9は、増幅されたアナログ音声信号をサンプリングレート16kHzでサンプリングして、16ビットのデータブロックを有し、かつオーディオ情報A1として送ることができるデジタルオーディオデータ形式で、サンプリングされた値をUSB接続手段12に送るように構成されている。

【0017】

このことは、第一アナログオーディオ信号A1がディクテーションマイクロフォンの中でデジタル化されることと、増幅段8を、ディクテーションマイクロフォン2の製造段階という早い段階においてマイクロフォン6の最大出力電圧に合わせて調整できるという利点を提供する。この方法では、増幅段8が過変調されたことに起因してボイス情報が失われることはなく、このためデジタルオーディオデータを含むデジタルオーディオ情報A1が良好な品質を有する。

【0018】

デジタルオーディオデータを含むオーディオ情報A1は、USB接続手段12によって、オーディオ処理手段7のデジタル-アナログコンバータ11に送ることができる。デジタルオーディオデータに対応するアナログオーディオ信号は、デジタル-

アナログコンバータ段11によって、第二増幅段10に送ることができる。第二増幅段10によって送られる第二アナログオーディオ信号AS2は、スピーカに与えて音響的に再現できる。

【0019】

ディクテーションマイクロフォン2は、さらに、制御情報S1を手動で入力するための入力手段14を含み、この入力手段は、キーボード15、トラックボール16、入力処理手段17を含む。キーボード15は、特に、録音キー18、停止キー19、順送りボタン20、早送り巻き戻しボタン21を有する。キーボード15は、キー情報TS1を入力処理手段17に送るはたらきをする。

【0020】

トラックボール16は、ディクテーションマイクロフォン2の穴の中に埋め込まれたボールによって形成され、このボールは、使用者が回転させて、例えば、モニタ4に表示されるカーソルを動かすことができる。トラックボール16を使用することで、トラックボール情報TB1を入力処理手段17に送ることができる。入力処理手段17は、そこに送られたキー情報TS1とトラックボール情報TB1を処理して、各制御情報S1をUSB接続手段12に送るように構成されている。

【0021】

USB接続手段12は、オーディオ処理手段7と入力手段14との両方に供給電圧Uを供給して、オーディオ処理手段7と入力手段14に含まれている段に電力を供給することができる。このことは、ディクテーションマイクロフォン2が、ディクテーションマイクロフォン2の段からの情報と供給電圧Uの両方を1本のケーブルリンクのみを介して受け取るという利点を提供する。

【0022】

USB接続手段12は、ディクテーションマイクロフォン2の第一端子22を介してコンピュータ1のUSB接続手段23に接続される。ディクテーションマイクロフォン2のUSB接続手段12は、デジタルデータバスリンクを設定するようにされ、さらに、利点として、コンピュータ1のUSB接続手段23へのUSBデータバスリンク(Universal Serial Bus: Specification version 1.0(1月15日)およびversion 2.0)を設定するように構成されている。

【0023】

USBデータバスリンク上で、USB接続手段12または23に送られたオーディオ情報AIのオーディオデータまたは制御情報SIの制御データに、誤り訂正コードが加えられ、これらのデータは一括して送信データUDとしてUSB接続手段23または12に送信される。送信されたデータUDが受信されるUSB接続手段12または23は、誤り訂正コードを評価し、必要に応じ、オーディオデータまたは制御データの中の、送信時に発生する誤りを訂正するように構成されている。USBデータバスリンクにおいて、制御情報SIの制御データは、データ送信速度1.6 MB/秒で送信され、オーディオ情報AIのオーディオデータは、データ送信速度12MB/秒で送信される。

【0024】

図3は、図1に示されているフットスイッチ3を、ブロック線図の形式で詳細に示す。フットスイッチ3は、制御情報SIを足で入力するための入力手段24を有する。入力手段24は、オーディオ再生スイッチ25、停止スイッチ26、入力処理手段27を含む。オーディオ再生スイッチ25または停止スイッチ26によって生成されたスイッチ情報SCIは、入力処理手段27に送ることができる。入力処理手段27は、そこに与えられたスイッチ情報SCIを処理して、各制御情報SIをフットスイッチ3のUSB接続手段28に与えるように構成されている。

【0025】

フットスイッチ3は、さらに、デジタル-アナログコンバータ段30と第三増幅段31を有するオーディオ処理手段29を含む。オーディオデータを含むオーディオ情報AIは、デジタル-アナログコンバータ段30にUSB接続手段28によって与えることができ、デジタル-アナログコンバータ段は、オーディオデータに対応するアナログオーディオ信号を第三増幅段31に送る。第三増幅段31によって送られる第三アナログオーディオ信号A3は、フットスイッチ3内に備えられたスピーカ32と第二端子33の両方に与えることができる。ヘッドフォン34は、第二端子33に接続されている。スピーカ32とヘッドフォン34によって、スピーチ伝達情報を音響的に再現することが可能であり、これについては後にさらに説明する。

【0026】

オーディオ処理手段29がフットスイッチ3の中に備えられているので、ヘッドフォン34またはスピーカをフットスイッチ3に直接接続できるという利点を得られる。結果として、ディクテーションマシンも形成するコンピュータ1は、オーディオ処理手段を有するサウンドカードを有する必要がない。

【0027】

フットスイッチ3のUSB接続手段28は、USB接続手段12と23に相当する。USB接続手段28は、入力処理手段24とオーディオ処理手段29の両方に供給電圧Uを供給して、入力処理手段24とオーディオ処理手段29に含まれている段に電力を供給できる。このことは、フットスイッチ3が伝達情報のみでなく供給電圧Uも1本のケーブルリンクのみを介して受け取るという利点を提供する。

【0028】

フットスイッチ3のUSB接続手段28は、コンピュータ1のUSB接続手段23に第三端子35を介して接続されている。コンピュータ1は、USB接続手段23とオーディオメモリ手段37を有する受信手段36を含み、話し手によって発音されたスピーチ伝達情報を受信するように構成されている。ディクテーションマイクロフォン2のマイクロフォン8に話し手によって発音されたスピーチ伝達情報は、デジタルオーディオデータによって構成されるオーディオ情報A1としてUSB接続手段12からUSB接続手段23に送信させ、また、USB接続手段23からオーディオメモリ手段37にオーディオ情報A1として送信させて格納させることができる。ディクテーションマイクロフォン2のUSB接続手段12またはフットスイッチ3のUSB接続手段28からコンピュータ1のUSB接続手段23に与えられる制御情報S1は、USB接続手段23によって受信して、コンピュータ1のさらなる手段に送ることができ、さらなる手段については後にさらに説明する。

【0029】

コンピュータ1によって形成されるスピーチ認識装置は、スピーチ係数インジケータSKIを格納するために備えられたスピーチ係数メモリ手段38を含む。スピーチ係数インジケータSKIは、音声モデルデータSM1、コンテキスト情報PI、音楽参照情報PRIを含む。

【0030】

コンテキスト情報PIは、コンテキストのメモリ手段39に格納され、スピーチ伝達情報の中の、スピーチ認識装置によって認識できるワードすべてと、スピーチ伝達情報の中のワードの発生確率を示すワード頻度メーターとを含む。音声モデルデータSMIは、音声モデルメモリ手段40に格納され、ワードシーケンスと、スピーチ伝達情報の中のこれらのワードシーケンスの発生の頻度を示すワードシーケンス頻度メーターとを含む。音素参照情報PRIは、音素参照メモリ手段41に格納され、話し手の発音の特性に合わせたスピーチ認識装置の調整を可能にする。

【0031】

スピーチ認識装置は、さらに、スピーチ認識手段42を含む。USB接続手段23は、スピーチ認識方法を実施するためにスピーチ認識モードをアクティブにするために、制御情報SIをスピーチ認識手段42に送ることができる。スピーチ認識手段42は、スピーチ認識モードがアクティブにされているときに、オーディオメモリ手段37に格納されているオーディオ情報AIを読み取り、スピーチ係数メモリ手段38に格納されているスピーチ係数インジケータSKIを読み取るように構成されている。

【0032】

スピーチ認識方法の実行中に、スピーチ認識手段42は、オーディオ情報AIとスピーチ係数インジケータSKIを評価するように構成されている。次いで、スピーチ認識手段42は、オーディオ情報AIの各パートについて、オーディオ情報AIの各パートに一致する複数の可能なワードシーケンスを決定し(一致するシーケンスがある場合)、オーディオ情報AIのすべてのパートについて、これらの可能なワードシーケンスを可能性のあるテキスト情報PTIとして生成する。可能性のあるテキスト情報PTIは、オーディオ情報AIの例えば、20個の可能なワードシーケンスを含む。さらに、スピーチ認識手段42は、スピーチ認識方法が実行されると、そのたびに、オーディオ情報AIの各パートについて、もっとも可能性の高いワードシーケンスを決定し、すべてのオーディオ情報AIの、これらのもっとも可能性の高いワードシーケンスを、認識されたテキスト情報RTIとして生成する。

【0033】

スピーチ認識装置は、テキスト情報を格納するためのテキストのメモリ手段43

を有する。テキストのメモリ手段43は、可能性のあるテキストのメモリ手段44、認識されたテキストのメモリ手段45、訂正されたテキストのメモリ手段46、トレーニングテキストのメモリ手段47を含む。スピーチ認識方法の実行中にスピーチ認識手段42によって生成される可能性のあるテキスト情報は、可能性のあるテキストのメモリ手段44に格納でき、認識されたテキスト情報RTIは、認識されたテキストのメモリ手段45に格納できる。

【0034】

コンピュータ1は、さらに、例えば、Word For Windows（登録商標）95などのテキスト処理プログラムを実行するように配置されており、かつ、テキスト処理手段48を形成する。スピーチ認識方法の実行中にスピーチ認識手段42によって生成された認識されたテキスト情報RTIは、テキスト処理手段48に送ることができる。テキスト処理手段48は、認識されたテキスト情報RTIを含むピクチャ情報PIをモニタ4に転送するように構成されている。

【0035】

ディクテーションマイクロフォン2の入力手段14によって、入力制御情報SIを入力でき、この制御情報SIは、スピーチ認識手段42とテキスト処理手段48に与えることができ、スピーチ認識装置の訂正モードをアクティブにする。テキスト処理手段48、モニタ4、キーボード5は、訂正手段49を構成する。訂正モードがアクティブであるとき、使用者は、キーボード5によって、認識されたテキスト情報RTIを訂正し、訂正されたテキスト情報OTIを訂正されたテキストのメモリ手段46に格納することができる。このことについては後にさらに説明する。

【0036】

トレーニングテキストのメモリ手段47には、特定の分野（ビジネスレター、学問分野：放射線医学、学問分野：法律など）の典型的なワードフレーズやワードシークエンスを含むトレーニングテキスト情報TTIが格納されている。制御情報SIは、ディクテーションマイクロフォン2の入力手段14によって入力でき、この制御情報は、スピーチ認識手段42とテキスト処理手段48に与えて、スピーチ認識装置の初期トレーニングモードをアクティブにできる。

【0037】

テキスト処理手段47は、初期トレーニングモードがアクティブのとき、トレーニングテキストのメモリ手段47からトレーニングテキスト情報TTIを読み出して、各ピクチャ情報PIをモニタ4に送るように構成されている。使用者は、これによってモニタ4に表示されたトレーニングテキストをマイクrophon6に話すことで、使用者の発音タイプに合わせてスピーチ認識装置を調整することができる。

【0038】

スピーチ認識装置は、スピーチ係数メモリ手段38に格納されているスピーチ係数インジケータSKIを、使用者の発音のタイプと、使用者によって普通に使われるワードおよびワードシーケンスに合わせて調整するための調整手段50を有する。テキストのメモリ手段43、訂正手段49、調整手段50は、まとめてトレーニング手段51を形成する。スピーチ係数インジケータSKIのこの調整は、使用者によって読まれるトレーニングテキスト情報TTIが知られている状態で、初期トレーニングモードがアクティブであるときに行われる。

【0039】

しかし、この調整は、ボイス情報に一致するテキスト情報が認識されたテキスト情報RTIとして認識され、かつ使用者によって訂正されたテキスト情報CTIに訂正される、調整モード時にも行われる。この目的のため、トレーニング手段51は、テキスト比較手段52を有し、このテキスト比較手段52は、認識されたテキスト情報RTIと訂正されたCTIを比較して、少なくとも1つの一致インジケータCIを決定するように構成されている。調整モードがオンのとき、テキスト比較手段52において、図4に示されている調整テーブル53が確立され、このテーブルについては後にさらに説明する。

【0040】

調整情報NIである調整の結果は、調整手段50によってスピーチ係数メモリ手段38に送ることができ、この調整情報は、格納されているスピーチ係数インジケータSKIを調整する。さらに、トレーニングインジケータTIは、トレーニング手段51の調整手段50によって送ることができ、このトレーニングインジケータTIは、スピーチ係数メモリ手段38に格納されているスピーチ係数インジケータSKIの調整の程度を示す。トレーニングインジケータTIは、スピーチ係数メモリ手段38に

格納されているスピーチ係数インジケータSKIが、使用者に合わせてすでに調整された回数と調整されたワード数についての情報を含む。スピーチ係数インジケータSKIが使用者に合わせて調整される回数が多いほど、その使用者にとってのスピーチ認識装置の認識率は良好になる。

【0041】

スピーチ認識装置は、スピーチ係数メモリ手段38内にスピーチ係数インジケータSKIをインポートするため、および/または、スピーチ係数メモリ手段38に格納されているスピーチ係数インジケータSKIをエクスポートすることを可能にする転送手段54を含む。この目的のため、使用者は、キーボード6によって転送情報TRIを入力して、それを転送手段54に送ることができ、これについては後にさらに説明する。

【0042】

第一応用例を参照しながら、ディクテーションマイクロフォン2、フラットスイッチ3、コンピュータによって形成されるディクテーションマシンとスピーチ認識装置とについて、さらに説明する。第一応用例では、コンピュータ1の使用者は、「Quality」タイプの3台のコンピュータを知人「John」に注文するための手紙をディクテーションしようとしていることを想定する。この目的のため、使用者は、ディクテーションマイクロフォン2の録音キー18を押し、「Dear John I herewith order three Quality computers for my office Harry」とディクテーションする。すると、このボイス情報に対応するオーディオ情報AIが、ディクテーションマイクロフォン2から送信データUDとしてUSB接続手段23に送信され、最終的にオーディオメモリ手段37に格納される。この処理の間、ディクテーションマイクロフォン2とコンピュータ1は、ディクテーションマシンを構成する。

【0043】

ディクテーションマイクロフォン2の録音キー18を作動させることによって、スピーチ認識手段42のスピーチ認識モードをアクティブにする制御情報SIが、ディクテーションマイクロフォン2によって、USB接続手段23を通じて送信データUDとしてスピーチ認識手段42に与えられる。スピーチ認識モードがアクティブにされたスピーチ認識手段42内において、オーディオメモリ手段37に格納されている



オーディオ情報AIが、読み出され、スピーチ係数メモリ手段38から読み取られたスピーチ係数インジケータSKIと共に評価される。次いで、スピーチ認識手段42は、複数のワードシーケンスを可能性のあるテキスト情報PTIとして決定し、これらを可能性のあるテキストのメモリ手段44に格納する。オーディオ情報AIの最初のパートに対して、可能性のあるテキスト情報PTIは、例えば、ワードシーケンス「Hear John why」、「Dear John I」、および「Dear John hi」を含むことが出来る。

【0044】

スピーチ認識手段42は、テキスト情報「Dear John I herewith organ three quality computers for my office Mary」を、もっとも可能性の高いテキスト情報として決定し、これを認識されたテキスト情報RTIとして認識されたテキストのメモリ手段46に送る。発声されたワード「order」はワード「organ」と誤って認識され、発声されたワード「Harry」はワード「Mary」と誤って認識された。

【0045】

スピーチ認識手段42によって認識され、かつ認識されたテキストのメモリ手段46に格納された認識されたテキスト情報RTIは、次いで、テキスト処理手段48によって読み出され、モニタ4に表示される。使用者は、自分が言った2つのワード「order」と「Harry」が誤って認識されたことを認識し、認識されたテキスト情報RTIを訂正するために、この使用者は、ディクテーションマイクロフォン2の入力手段14によってスピーチ認識装置の訂正モードをアクティブにする。

【0046】

訂正モードがアクティブにされると、使用者は、フットスイッチ3のオーディオ再生スイッチ25を押すことで、オーディオメモリ手段37に格納されているオーディオ情報AIを、ヘッドフォンまたはスピーカ32によって音響的に再生することができ、その一方で、オーディオ情報AIから認識されたテキスト情報は、再生されるオーディオ情報AIと同期してモニタ4に表示される。このような同期的な編集または訂正は、以前から知られている。

【0047】

認識されたテキスト情報RTIの使用者による第一タイプの訂正において、使用

者は、誤って認識されたワード「forgan」と「Mary」のみを、最初に発声された2つのワード「forder」と「Harry」に訂正する。訂正されたテキスト情報CTIが、テキスト処理手段48内に構築され、訂正されたテキストのメモリ手段46に格納される。最初の訂正ステップのあとに、訂正されたテキストのメモリ手段46に格納されたこの訂正されたテキスト情報CTIは、スピーチ係数インジケータSKIを調整するのに非常に適している。以降のディクテーションにおいて使用者によって発声されるワード「forder」と「Harry」は、すでに調整されたこのスピーチ係数インジケータSKIによって、次のスピーチ認識処理においては正しく認識される。

【0048】

経験的には、使用者は、訂正モードがアクティブのときに、第二タイプの別の訂正も行う。第一応用例では、使用者は、「Standard」タイプのコンピュータ8台も注文したいことを思い出し、このため使用者は、すでに訂正されたテキスト情報CTIに、キーボード5によって、ワードシーケンス「and three Standard computers」を挿入する。さらに、使用者は、知人「John」宛の手紙の住所が含まれているテキストAを、すでに訂正されたテキスト情報CTIの先頭の前にコピーする。最後に、使用者は、このような注文の通例として合意されている割引率についての脚注を含むテキストBを、訂正されたテキスト情報CTIの中のすでに訂正されたテキスト情報CTIの後にコピーする。

【0049】

次いで、使用者は、ディクテーション時に使用者によって発声されたボイス情報とまったく関係ないテキスト部分を、認識されたテキスト情報RTIに追加する。このようなテキスト部分は、音素参照情報PRIを調整するために評価されてはならない。何故ならば、対応するオーディオ情報AIが、オーディオメモリ手段97に格納されていないためである。さらに、このようなテキスト部分を使って、コンテキスト情報Wまたは音声モデルデータSMIを調整することは無益である。何故ならば、テキストAとテキストBは、使用者の訂正されたテキスト情報CTIに頻繁に含まれるが、ディクテーション時に使用者によって発声されることは殆どないためである。

【0050】

ボイス情報とまったく関係ないテキストパートを使ってコンテキスト情報W1と音声モデルデータSM1がトレーニングされると、例えば、知人Johnの住所のテキストAに含まれるストリート名「Hell Road」というワードシーケンスは、高い発生確率を有する。何故ならば、使用者は、Johnに頻繁に手紙を書き、訂正されたテキスト情報CT1に必ず住所をコピーするからである。しかし、ディクテーション時に発声されるワードシーケンス「Hello Rod」は、容易に誤って認識されてしまうであろう。何故ならば、ワードシーケンス「Hello Rod」は使用者のディクテーションにおいて相対的に頻繁に発生するにもかかわらず、音声モデルメモリ手段40におけるワードシーケンス「Hello Rod」の発生の確率は、ワードシーケンス「Hell Road」に比べて低くなるためである。

【0051】

使用者が、必要な訂正をすべて行い、訂正されたテキスト情報CT1が訂正されたテキストのメモリ手段46に格納された後、使用者は、ディクテーションマイクロフォン2のキーボードのキーを押して、スピーチ認識装置の調整モードをアクティブにして、スピーチ認識装置のスピーチ係数インジケータSK1を調整する。このとき、ディクテーションマイクロフォン2の各制御情報81は、USB接続手段23を介してスピーチ認識手段42に与えられる。

【0052】

調整モードがアクティブにされたスピーチ認識装置において、調整手段50は、オーディオメモリ手段37に格納されているオーディオ情報A1、つまりテキストのメモリ手段43に格納されている可能性のあるテキスト情報PT1や認識されたテキスト情報RT1、訂正されたテキスト情報CT1などを読み取るように構成されている。さらに、アクティブにされた調整モードにおいて、テキスト比較手段52は、スピーチ係数インジケータSK1を調整するのに適した、訂正されたテキスト情報CT1のテキストパートを決定するために、図4に示されている調整テーブル53を用意するように構成されている。

【0053】

この目的のため、まず、認識されたテキスト情報RT1と訂正されたテキスト情

報CTIが調整テーブル53に書き込まれ、その一方で、誤って認識されたワードと、訂正もしくは挿入されたワードとテキストパートが太字で表される。テキスト比較手段52は、認識されたテキスト情報RTIと訂正されたテキスト情報CTIを比較し、一致するワードを調整テーブルの各位置における灰色のフィールドによって表すように構成されている。認識されたテキスト情報RTIのワード群「three computers」を有するテキストパートは、訂正されたテキスト情報CTIの中に2回見出される。

【0054】

さらに、テキスト比較手段52は、認識されたテキスト情報RTIと訂正されたテキスト情報CTIを比較するときに、各テキストパートの一致インジケータCIを決定するように構成されている。このため、テキスト比較手段52は、テキストパートに含まれる、灰色のフィールドによって表された一致ワードの数を調べる。さらに、テキスト比較手段52は、訂正されたテキスト情報CTIにおけるワードの挿入、削除、置換それぞれに対して1減点が与えられるものとして、各テキストパートの減点を決定する。テキストパートの一致インジケータCIは、テキストパートの中の一致ワードの数と減点の数から決定される。

【0055】

テキスト比較手段52において、一致インジケータCIの最小値MWが決定される。各テキストパートにおいて、訂正されたテキスト情報CTIの中の隣接するワードの訂正に対する減点が4個以上与えられたときには、この最小値に達しないものとする。スピーチ係数インジケータSKIの調整には、一致インジケータCIが最小値MWを超えるテキストパートのみが使われる。

【0056】

テキスト比較手段52は、調整テーブル53を評価し、訂正されたテキスト情報CTIにテキストAのワードすべてとテキストBのワードすべてが挿入されたものと認識する。何故ならば、テキストAのワード数とテキストBのワード数に一致する数の減点が、テキスト比較手段52によってテキストAとテキストBに与えられるからである。従って、テキスト比較手段52は、テキストAとテキストBに対し最小値MWより小さい値の一致インジケータCIを調整手段50に与える。このことは、ボイス

情報とまったく関係ない、訂正されたテキスト情報CTIのテキストパートが、スピーチ係数インジケータSKIの調整に使われないという利点を提供する。

【0067】

調整テーブル53によると、スピーチ係数インジケータSKIの調整に使うことができるテキストパートは、訂正されたテキスト情報CTIの経路P1と経路P2の可能性を有する。

【0068】

経路P1は、最初のテキストパート「Dear John I herewith」を含み、このパートは、テキスト比較手段52から減点がまったく与えられず、このためその一致インジケータは最小値MWを超える。経路P1は、最初のテキストパートの後に2番目のテキストパート「order three Quality computers and」を含み、この2番目のパートは、ワード「order three Quality computers」の挿入と、ワード「and」のワード「organ」への置換に対して、5個の減点(1)～(5)が与えられていて、このためその一致インジケータは最小値MWに満たない。経路P1は、2番目のテキストパートの後に3番目のテキストパート「three Standard computers for my office Harry」を含み、このパートは、ワード「Standard」の挿入に対して1つの減点(6)しか与えられていないので、その一致インジケータCIは最小値MWを超える。従って、経路P1に含まれる第一テキストパートと第三テキストパートが、スピーチ係数インジケータSKIの調整に原理的には適している。

【0069】

経路P2は、4番目のテキストパート「Dear John I herewith order three Quality computers」を含み、このパートは、ワード「organ」をワード「order」に置き換えたことに対して、テキスト比較手段52から1個しか減点(1)が与えられていないので、その一致インジケータCIは最小値MWを超える。経路P2は、4番目のテキストパートのあとに5番目のテキストパート「and three Standard computer s」を含み、このパートは、これらのワードの挿入に対して4個の減点(2)～(5)が与えられているので、その一致インジケータは最小値MWに満たない。経路P2は、5番目のテキストパートのあとに6番目のテキストパート「for my office Harry」を含み、このパートは、ワード「Mary」を「Harry」に置き換えたことに対し

て1個の減点(6)が与えられているので、その一致インジケータCIは最小値MWを超える。従って、経路P2に含まれる第4テキストパートと第6テキストパートが、スピーチ係数インジケータSKIの調整に原理的には適している。

【0060】

テキスト比較手段52は、合計一致インジケータSCIが最大値を有する経路P1またはP2のテキストパートを決定するように構成されている。この目的のため、テキスト比較手段52は、訂正されたテキスト情報RTIの経路P1の8個のワードと経路P2の10個のワードが、認識されたテキスト情報RTIのワードに完全に一致し、調整テーブルにおいて灰色のフィールドとして表されていることを判断する。さらに、テキスト比較手段52は、各経路の減点を合計し、経路P1については7個の減点、経路P2については6個の減点と決定する。

【0061】

最終的に、テキスト比較手段52は、10個の一致ワードと6個のみの減点を有する経路P2のテキスト情報が、認識されたテキスト情報RTIとのより大きな一致程度を有すると判断し、経路P2のテキストパート群に対してそれぞれより高い合計一致インジケータSCIを調整手段50に供給する。

【0062】

このことは、スピーチ係数インジケータSKIの調整において、第一タイプの訂正によって訂正されたワード(例えば、「order」や「Harry」)が調整に使われ、第二タイプの訂正によって訂正され、ボイス情報とまったく関係ないテキストパート(例えば、「and three Standard computers」)がスピーチ係数インジケータSKIの調整に使われないという利点を提供する。この結果、スピーチ係数インジケータSKIは、使用者の発音の特性と使用者の典型的なワード形成に合わせて非常に迅速かつ良好に調整され、スピーチ認識装置の特に良好な認識率が達成される。

【0063】

調整手段50は、調整モードがアクティブにされたスピーチ認識装置において、テキスト比較手段52から合計一致インジケータSCIを受信した後、訂正されたテキストのメモリ手段48からの訂正されたテキスト情報CTIの経路P2のテキストパ

ートを読み取って、これらのテキストパートを評価するように構成されている。

【0064】

コンテキスト情報PIを調整するために、ワード「order」P2の4番目と8番目のテキストパートの各ワード「Dear」、「John」、「I」、...、「Harry」には、コンテキストのメモリ手段39にそれぞれ格納されている頻度メーター（この頻度メーターは、使用者のディクテーションにおける各ワードの発生の確率を表す）で値「1」が加えられる。調整に使われるテキストパートのワードが使用者によって初めてディクテーションされたときには、そのワードは、頻度メーターの値「1」とともにコンテキストのメモリ手段39に格納される。

【0065】

音声モデルメモリに格納されている音声モデルデータSMIの調整のため、例えば、ワードシーケンス「I herewith order」に割り当てられた格納されている頻度メーターは、使用者のディクテーションにおいてこのようなワード形成の発生確率がより高いことを表すために、値「1」だけ大きくされる。

【0066】

音素参照メモリ手段41に格納された音素参照情報PRIの調整のため、ワードシーケンスP2の4番目と8番目のテキストパートのワードのみが、認識されたテキスト情報RTIのワードに完全に一致かつ調整テーブル53で灰色のフィールドとして表されている対応オーディオ情報AIとの組み合わせで、評価される。さらに、対応するオーディオ情報AIをもつ可能性のあるテキスト情報PTIのワードも評価することができる。この目的のため、調整手段50は、第一タイプの訂正によって使用者によって置換されたワード（「order」、「Harry」）が、オーディオ情報AIのこのパートに対して、可能性のあるテキストのメモリ手段44に格納されているか否かを調べる。このようなテキストが、可能性のあるテキストのメモリ手段44の中に見出されると、対応するオーディオ情報AIをもつそのワードも、音素参照情報PRIの調整のために評価される。

【0067】

次いで、これらの調整を含む調整可能な情報NIは、スピーチ係数インジケータSKIを調整するために、調整手段50によってスピーチ係数メモリ手段38に与えら

れる。同様に、調整手段50は、スピーチ係数メモリ手段38にすでに格納されているスピーチ係数インジケータSKIがすでにトレーニングされたワードの数を示すトレーニングインジケータTIを、転送手段54に与える。このことは、スピーチ認識装置が使用者に合わせてすでに良好に調整されたか否かと、それによってスピーチ認識装置がその使用者にとっておそらく良好な認識率を有するか否かが、トレーニングインジケータTIによってただちに明らかになるという利点を提供する。

【0068】

第二応用例によると、コンピュータ1によって形成されるスピーチ認識装置の使用者が、第二スピーチ認識装置を形成する第二コンピュータをさらに有すると想定されている。この場合、使用者が、コンピュータ1によって形成されるスピーチ認識装置(そのスピーチ認識情報は数ヶ月にわたって使用者によってトレーニングされている)で達成された認識率と同等に良好な認識率を第二スピーチ認識装置上でも達成することを望むことは明らかである。

【0069】

この目的のため、使用者はキーボード5のキーを作動させ、キーボード5によって転送情報TRIを転送手段54に送る。転送手段54は、スピーチ係数メモリ手段38に格納されているスピーチ係数インジケータSKIを読み取る。転送手段54は、さらに、読み取られたスピーチ係数インジケータSKIとスピーチ係数インジケータSKIのトレーニングインジケータTIを、コンピュータ1のディスクドライブに挿入されたディスク55に格納する。

【0070】

このことは、スピーチ係数メモリ手段38に格納され、かつ使用者によってすでにトレーニングされたスピーチ係数インジケータSKIが、コンピュータ1からエクスポートできるというメリットを提供する。さらに、スピーチ係数インジケータSKIは、使用者の第二コンピュータにインポートできる。結果として、使用者は、さらに、使用者によってトレーニングされたスピーチ係数インジケータSKIを、任意の台数の別のコンピュータでも使用でき、それらのコンピュータは、図1に示されているスピーチ認識装置に相当するスピーチ認識装置を形成する。この



方法により、使用者は、複数のコンピュータにおいてもスピーチ認識装置の良好な認識率を得る。

【0071】

スピーチ係数インジケータSKIがエクスポートされる時には、スピーチ係数インジケータSKIのトレーニングインジケータTIもエクスポートされる。一方、転送手段54からスピーチ係数インジケータSKIがインポートされる時には、スピーチ係数メモリ手段38に格納済みのスピーチ係数インジケータSKIよりも、インポートするスピーチ係数インジケータSKIの方がより良好にトレーニングされているかを、さしあたり確認できる。転送手段54は、インポートしようとするスピーチ係数インジケータSKIが、スピーチ係数メモリ手段38に格納されているスピーチ係数インジケータSKIよりもすでに良好にトレーニングされていることがテストによって証明された場合にのみ、そのスピーチ係数インジケータSKIをインポートする。

【0072】

第三応用例によると、コンピュータ1が、インターネットに接続されたインターネットサーバを形成し、第4端子によってインターネットに接続されていると想定されている。さらに、ディクテーションマシンを形成するさらなるコンピュータの使用が、コンピュータ1のホームページメモリ手段57に格納されたインターネットサーバのホームページからテキスト情報を照会し、そのテキスト情報が第4端子56を介してそのさらなるコンピュータに送られたものと想定する。ホームページからのこのテキスト情報は、通常行われているように、ホームページメモリ手段57に格納されているさらなるテキスト情報を照会することができるリンク情報を含む。図1に示されているインターネットサーバのスピーチ認識装置を使用することにより、さらなるコンピュータの使用が、ホームページ上に示されているリンク情報を、さらなるコンピュータによって送られるボイス情報によってアクティブにすることができる。

【0073】

第三応用例によると、さらに、使用者が、ホームページに表示されているリンク情報「hotels」をアクティブにすることを望んでいるとを想定する。この目的

のため、使用者は、さらなるコンピュータのディクテーションマイクロフォンの録音キーを作動させ、ディクテーションマイクロフォンのマイクロフォンにボイス情報「hotel」を発音する。このボイス情報に含まれるオーディオ情報AIは、第4端子56を介してコンピュータ1に与えられ、オーディオ情報AIがオーディオメモリ手段37に格納される。このボイス情報によってアクティブにされたコンピュータ1のスピーチ認識装置のスピーチ認識モードにおいて、ワード「hotel」が、スピーチ認識手段42によって、認識されたテキスト情報RTIとして認識され、このリンク情報に割り当てられたテキスト情報が、第4端子56を介してホームページメモリ手段57から使用者のさらなるコンピュータに送られる。

【0074】

スピーチ認識モードがアクティブにされる前に、さらなるコンピュータによってエクスポートされ、かつインターネットを介して第4端子56に与えられたスピーチ係数インジケータSKIは、有利に、転送手段54によってコンピュータ1にインポートされる。この結果、コンピュータ1によって形成されるインターネットサーバは、各使用者に合わせて調整され、これにより、きわめて高い認識率が達成される。

【0075】

使用者のスピーチ係数インジケータSKIは、インターネットサーバに送信される各オーディオ情報AIと共に送信できることも理解されるであろう。しかし、インターネットサーバのホームページが最初に照会されたときのみ、使用者のスピーチ係数インジケータSKIをインターネットサーバに送信して、その使用者のためにスピーチ係数インジケータSKIをインターネットサーバに格納しておくことは、特に有利である。それ以降、その使用者のさらなるコンピュータからオーディオ情報AIを受け取るたびに、インターネットサーバは、すでに格納されているスピーチ係数インジケータSKIをスピーチ認識用に使うことができる。

【0076】

訂正されたテキスト情報CTIの格納後、スピーチ係数インジケータSKIを調整するための調整モードを自動的にアクティブにできること、またはコンピュータ1のキーボード5のキーを介して使用者によってアクティブにできることも理解さ

れるであろう。

【0077】

経路のテキストパートの一致値、または経路のテキストパートの合計一致値を決定する方法として、実施例を参照しながら説明した以外の方法も可能であることも理解されるであろう。さらに、3回の連続する減点が与えられたときに最小値MWに達しないと結論することは、一例にすぎないものと理解すべきである。同様に、最小値MWに達しない条件として、テキストパートの特定のワード数における特定の減点数を定義することもできる。

【0078】

スピーチ係数インジケータSKIを調整するために原理的には適しているテキストパートの特定のセクションのみを、スピーチ係数インジケータSKIの調整に使用することも有利であることも理解されるであろう。この場合にテキストパートは、スピーチ係数インジケータSKIを調整するための第一実施例を参照しながら説明した方法によって決定される。例えば、テキスト比較手段52によって減点が与えられているワードと、調整テーブル53内でそのワードに隣接する、灰色のフィールドによって表されているワードのみが使われる。

【0079】

このことは、スピーチ認識処理中に誤って認識されて、使用者によって訂正（置換）されたワード（「order」、「Harry」）と、これらのワードに隣接するワードとが調整に使われるという利点を提供する。この場合に、テキストパートの訂正されたテキスト情報のうち、認識されたテキスト情報RTIに完全に一致し、かつ原理的に調整に適したセクションは、使用者によってまったく訂正されていない可能性があるため、調整に使われない。

【0080】

手に持つことができる入力装置の入力手段が、コンピュータに制御情報を与えて、コンピュータをパワーセーブモードからアクティブモードにするように配置されてもよいことも理解されるであろう。

【0081】

コンピュータにおける動作モードをアクティブにするために、足で操作する入

力する装置も一般に使用できることも理解されるであろう。コンピュータは、テレビ受像機を形成でき、テレビのチャンネルは、フットスイッチで選択でき、そのチャンネルにテレビ放送が受信され、モニタに表示される。

【0082】

入力装置のUSB接続手段を接続できるマイクロフォンまたはスピーカも、足で操作する入力装置に接続でき、その場合にマイクロフォンまたはスピーカはUSB接続手段も有することも理解されるであろう。この場合、足で操作する入力装置は、オーディオ処理手段を有する必要はなく、その理由は、マイクロフォンまたはスピーカ内のオーディオ処理手段がアナログオーディオ信号を送るように配置されるためである。

【0083】

デジタルデータバスリンクを、例えば、Bluetooth仕様(Bluetooth Special Interest Group, May 1998)によるリンクなどの無線周波数リンクによっても形成できることも理解されるであろう。このことは、入力装置とコンピュータの間にケーブルリンクが必要ないという利点を提供する。

【0084】

フットスイッチ3の第二端子33にスピーカも接続できることも理解されるであろう。

【図面の簡単な説明】

【図1】 手に持つことができる入力ユニットと足で作動させることができる入力ユニットが接続され、かつ、認識されたテキスト情報と訂正されたテキスト情報とを比較するためのテキスト比較手段を含み、かつ、スピーチ係数インジケータをインポートおよびエクスポートするための転送手段を含む、スピーチ認識装置をブロック線図の形式で示す。

【図2】 アナログオーディオ信号を処理してデジタルオーディオータを送るためのオーディオ処理手段を有する、図1に示されている手に持つことができる入力ユニットを、ブロック線図の形式で示す。

【図3】 ヘッドフォンを接続でき、かつ内蔵スピーカを含む、図1に示されている足で作動させることができる入力ユニットを、ブロック線図の形式で示す。

す。

【図4】 スピーチ認識装置のテキスト比較手段によって確立される調整テーブルであり、認識されたテキスト情報と訂正されたテキスト情報が入力されたテーブルを示す。

【符号の説明】

- 1 コンピュータ
- 2 ディクテーションマイクロフォン
- 3 フットスイッチ
- 4 モニタ
- 5 キーボード
- 6 マイクロフォン
- 7 オーディオ処理手段
- 8 第一増幅段
- 9 アナログ-デジタルコンバータ段
- 10 第二増幅段
- 11 デジタル-アナログコンバータ段
- 12 USB接続手段
- 13 スピーカ
- 14 入力手段
- 15 キーボード
- 16 トラックボール
- 17 入力処理手段
- 18 録音キー
- 19 停止キー
- 20 早送りボタン
- 21 巻き戻しボタン
- 22 第一端子
- 23 USB接続手段
- 24 入力手段

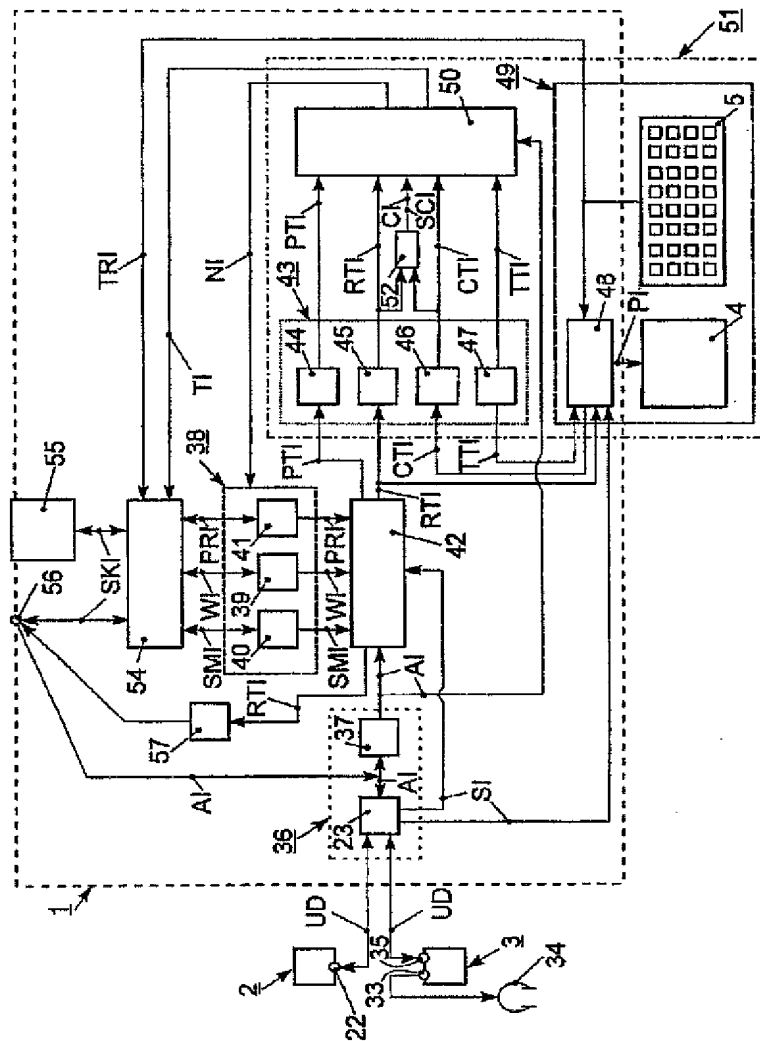
- 25 オーディオ再生スイッチ
- 26 停止スイッチ
- 27 入力処理手段
- 28 USB接続手段
- 29 オーディオ処理手段
- 30 デジタル-アナログコンバータ段
- 31 第三増幅段
- 32 スピーカ
- 33 第二端子
- 34 ヘッドフォン
- 35 第三端子
- 36 受信手段
- 37 オーディオメモリ手段
- 38 スピーチ係数メモリ手段
- 39 コンテキストのメモリ手段
- 40 音声モデルメモリ手段
- 41 音素参照メモリ手段
- 42 スピーチ認識手段
- 43 テキストのメモリ手段
- 44 可能性のあるテキストのメモリ手段
- 45 認識されたテキストのメモリ手段
- 46 訂正されたテキストのメモリ手段
- 47 トレーニングテキストのメモリ手段
- 48 テキスト処理手段
- 49 訂正手段
- 50 調整手段
- 51 トレーニング手段
- 52 テキスト比較手段
- 53 調整テーブル

S4 転送手段  
S5 ディスク  
S6 第4端子  
S7 ホームページメモリ手段  
AS1 第一アナログオーディオ信号  
AS2 第二アナログオーディオ信号  
AS3 第三アナログオーディオ信号  
AI オーディオ情報  
SI 制御情報  
SCI スイッチ情報  
TSI キー情報  
TBI トラックボール情報  
U 供給電圧  
SKI スピーチ係数インジケータ  
SMI 音声モデルデータ  
WI コンテキスト情報  
PRI 音楽参照情報  
PTI 可能性のあるテキスト情報  
CTI 訂正されたテキスト情報  
RTI 認識されたテキスト情報  
TTI トレーニングテキスト情報  
TI トレーニングインジケータ  
TRI 転送情報  
NI 調整情報  
PI ピクチャ情報  
MW 最小値  
SCI 合計一致インジケータ  
P1 経路  
P2 経路

37-80

37-30

【図 1】

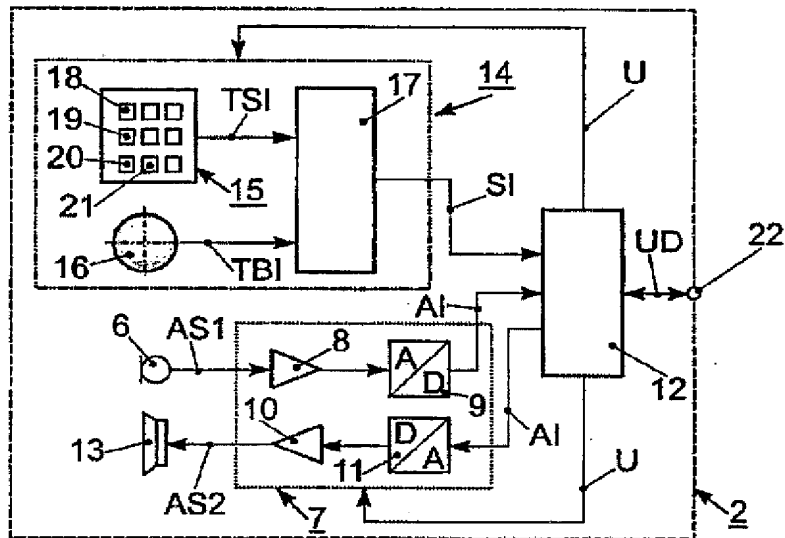


37-31

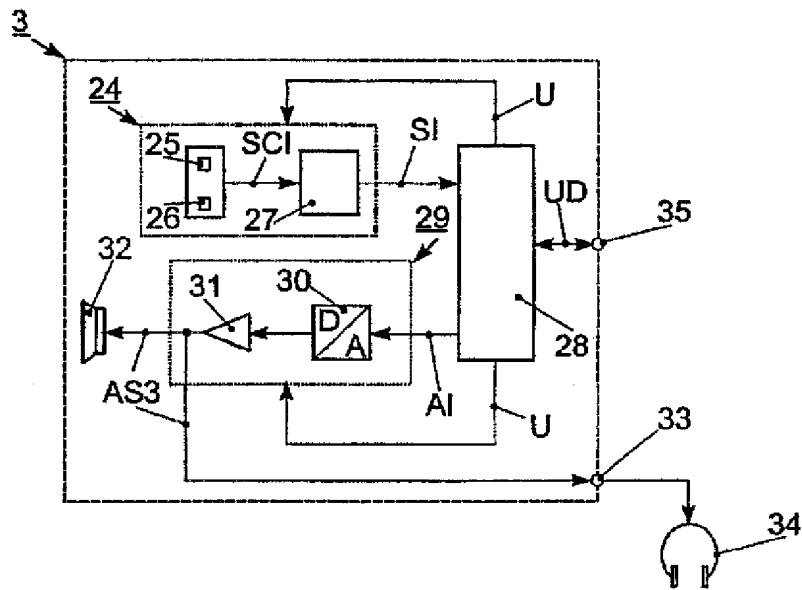
37-31



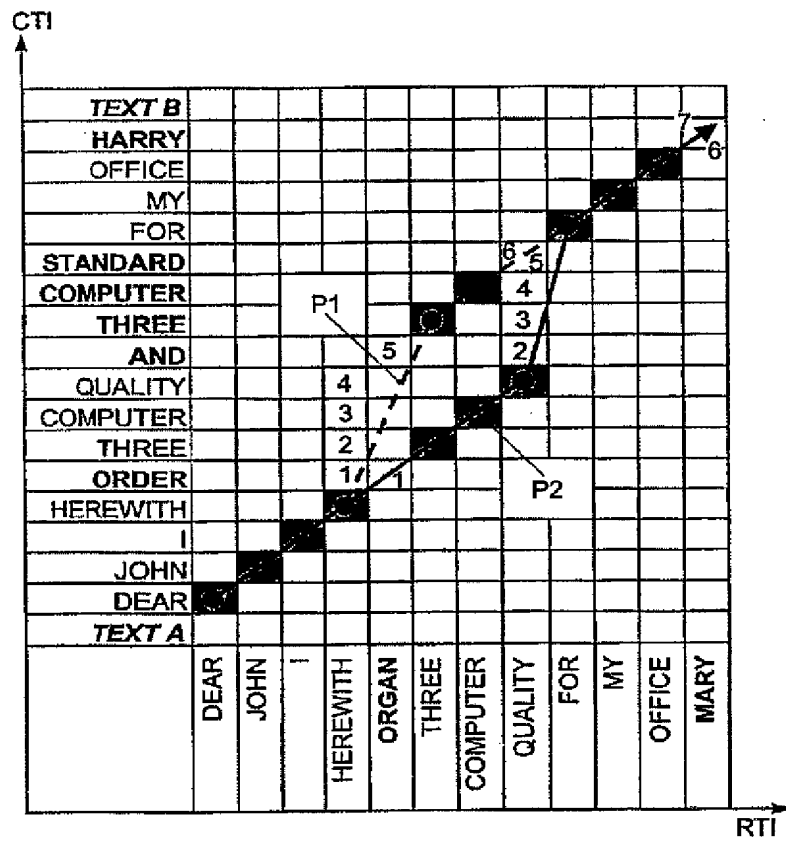
【図2】



【図3】



【图 4】



## INTERNATIONAL SEARCH REPORT

 Internat. Application No.  
PCT/JP 00/06167

<b>4. CLASSIFICATION OF SUBJECT MATTER</b> IPC 7 610L5/06		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>5. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) IPC 7 610L		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPD-Internal, WPI Data, ISM-TDS, EINSPEC		
<b>6. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category *	Citation of document, with indication, where appropriate, of the relevant passages Relevant to claim No.	
X	US 5 857 099 A (MITCHELL ET AL) 5 January 1999 (1999-01-05) abstract; figures 5, 8A-B, 9A-B column 1, line 39-48 column 2, line 64 -column 3, line 4 column 3, line 22-28	1-7
A	EP 0 773 532 A (DRAGON SYSTEMS INC) 14 May 1997 (1997-05-14) abstract page 2, line 22-28 page 3, line 1-14	1-3, 6, 7
A	US 5 852 001 A (HON ET AL) 22 December 1998 (1998-12-22) abstract; figures 2, 3 column 2, line 18-40	3-5
-/-		
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document not published on or after the international filing date "L" document which may raise doubts on novelty claim(s) or which is cited to establish the publication date of another claim or other specific reason (as specified) "D" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "C" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "I" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. "G" document member of the same patent family		
Date of the actual completion of the international search 4 October 2000		Date of mailing of the international search report 11/10/2000
Name and mailing address of the ISA European Patent Office, P.O. Box 5010 Patankou 2 NL - 3720 XZ Bilthoven Tel: (+31-70) 340-2043, Tlx 51 651 epe nl, Fax: (+31-70) 340-5018		Authorized officer Quéllavoine, R

Form PCT/ISA/210 (second page) July 2002

page 1 of 2

## INTERNATIONAL SEARCH REPORT

 PCT/JP 90/06167  
 PCT/JP 90/06167

C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim no.
A	<p>"ON-LINE DISTINCTION BETWEEN TEXT EDITING AND SPEECH RECOGNITION ADAPTION"</p> <p>IBM TECHNICAL DISCLOSURE BULLETIN, US, IBM CORP. NEW YORK,</p> <p>vol. 37, no. 10,</p> <p>1 October 1994 (1994-10-01), page 403</p> <p>XP000475716</p> <p>ISSN: 0018-8689</p> <p>the whole document</p>	1,6
A	<p>"DISTINCTION AT EXIT BETWEEN TEXT EDITING AND SPEECH RECOGNITION ADAPTION"</p> <p>IBM TECHNICAL DISCLOSURE BULLETIN, US, IBM CORP. NEW YORK,</p> <p>vol. 37, no. 10,</p> <p>1 October 1994 (1994-10-01), page 391</p> <p>XP000475710</p> <p>ISSN: 0018-8689</p> <p>the whole document</p>	1,6
A	<p>"ADDRESS THROUGH CORRECTION OF RECOGNITION ERRORS"</p> <p>IBM TECHNICAL DISCLOSURE BULLETIN, US, IBM CORP. NEW YORK,</p> <p>vol. 35, no. 1A, 1 June 1992 (1992-06-01), pages 226-227, XP000308844</p> <p>ISSN: 0018-8689</p> <p>the whole document</p>	3-5
A	<p>"SELECTION SENSITIVE MENUS FOR CORRECTION AND TRANSFER OF DICTATED TEXT"</p> <p>IBM TECHNICAL DISCLOSURE BULLETIN, US, IBM CORP. NEW YORK,</p> <p>vol. 37, no. 10,</p> <p>1 October 1994 (1994-10-01), page 365</p> <p>XP000475698</p> <p>ISSN: 0018-8689</p> <p>the whole document</p>	1,6

Form PCT/ISA/210 (Continuation of PCT/ISA/210) (July 1989)

page 2 of 2

37-35

37-35

# INTERNATIONAL SEARCH REPORT

Continuation of patent family members

International Application No.  
PCT/EP 00/06167

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5857099 A	05-01-1999	US 5799273 A	25-08-1998
		AU 4313897 A	17-04-1998
		GB 2302199 A, B	08-01-1997
		GB 2303955 A, B	05-03-1997
		WO 9813754 A	02-04-1998
EP 0773532 A	14-05-1997	US 5794189 A	11-08-1998
US 5852801 A	22-12-1998	NONE	

From PCT/EP/2000/0006167 (PCT/EP/2000/0006167)

---

フロントページの続き

(72)発明者 ミューラ ウォルタ  
オランダ国 5856 アー アー アインド  
ーフェン プロフホルストラーン 6

(72)発明者 シャッツ マーティン  
オランダ国 5856 アー アー アインド  
ーフェン プロフホルストラーン 6

Fターム(参考) 5D015 AA05 GG01 KK02 KK03 LL05

【要約の続き】

キスト情報(CTI)を比較し、少なくとも1つの一致インジ  
ケータ(CI)を決定するために備えられていて、調整手段  
(50)が、一致インジケータ(CI)が最小値(MW)を有する訂  
正されたテキスト情報(CTI)の1つのテキストパート(P2)  
のみを評価することによって、格納されたスピーチ係数  
インジケータ(SKI、PRI、SMI、WI)を調整するように構  
成されている、スピーチ認識装置。